



ESX ORF DNA Sequence (1 to 1116) → 1-phase Translation
DNA sequence 1116 b.p. ATGGCTGCAACC ... agtcggaaactgaa linear

371 Amino acids
MW: 41428 Dalton

1/1

ATG GCT GCA ACC TGT GAG ATT AGC AAC TAC TTC AGT GCG ATG TAC AGC TCG GAG GAC TCC ACC CTG GCC TCT GTT CCC
Met ala ala thr cys glu 1le ser asn 1le phe ser asn tyr phe ser ala met tyr ser ser glu asp ser thr leu ala ser val pro
91/31 121/41 151/51
CCT GCT GCC ACC TTT GGG GCC GAT GAC TTG GTC ACC CTG AGC AAC CCC CAG ATG TCA TTG GAG GGT ACA GAG GGC AGC TGG TTG
pro ala ala thr phe gly ala asp asp 1eu val leu ser asn pro gln met ser leu glu gly thr glu lys ala ser trp 1eu
181/61 211/71 241/81
GGG GAA CAG CCC CAG TTC TGG TCG ACG CAG CAG GAT GAC TGG ATC AGC TAC CAA GTG GAG AAG AAC TAC GAC GCA AGC GCC ATT
gly glu gln pro gln phe trp ser lys thr gln val 1eu asp trp 1le ser tyr gln val 1eu glu lys tyr asp ala ser ala ile
271/91 301/101 331/111
GAC TTC TCA CGA TGT GAC ATG GAT GCC ACC CTC TGC AAT TGT GCC CTT GAG GAG CTG CGT CTG GTC TTT GGG CCT CTG GGG GAC CAA
asp phe ser arg cys asp met asp gly ala thr leu cys asn cys ala leu glu glu 1eu arg leu val phe gly pro 1eu gly asp gln
361/121 391/131 421/141
CTC CAT GCC CAG CTG CGA GAC CTC ACT TCC AGC TCT TCT GAT GAG CTC AGT TGG ATC ATT GAG CTG CTG GAG AAG GAT GCC ATG GCC TTC
Iau his ala gln 1eu arg asp 1eu thr ser ser ser asp glu 1eu ser trp 1le 1le glu 1eu 1eu glu lys asp gly met ala phe
451/151 481/161 511/171
CAG GAG GCC CTA GAC CCA GGG CCC TTT GAC CAG GGC AGC CCC TTT GCC CAG GAG CTG GAC GGT CAG CAA GCC AGC CCC TAC CAC
gln glu ala 1eu asp pro gly pro phe asp gln gly ser pro phe ala gln gln gln ala ser pro tyr his
541/181 571/191 601/201
CCC GGC AGC TGT GGC GCA GGA GGC CCC TCC CCT GGC AGC TCT GAC GTC TCC ACC GCA GGG ACT GGT GCT TCT CGG AGC TCC CAC TCA
pro gly ser cys gly ala pro ser pro gly ser ser asp val ser thr ala gly thr gly ala ser arg ser ser his ser ser
631/211 661/221 691/231
GAC TCC GGT GGA AGT GAC GTG GAC CTG GAT CCC ACT GAT GGC AAG CTC TTC CCC AGC GAT GGT CGT TTT CGT GAC AAG AAG GGG GAT CCC
asp ser gly gly ser asp val 1eu asp pro thr asp gly lys 1eu phe pro ser asp gly phe arg asp cys lys 1ys gly asp pro

1/21

Fig. 1

2/21

811/271 AGA GGC ACC CAC CTG TGG GAG TTC ATC CGG GAC ATC CTC ATC AAC GAG CTC ACG GAG CTC ATG AAG TGG GAG AAT CGG CAT GAA
arg gly thr his 1eu trp glu phe 1le arg asp 1le 1eu flle his pro glu 1eu asn glu gly 1eu met lys trp glu asn arg his glu
901/301 931/311 961/321
66C GTC TTC TCC GGC TCC GAG GCT GTG GCC CAA CTA TGG GGC CAA AAG AAA AAG AAC AGC AAC ATG ACC TAC GAG AAG CTG AGC
gly val phe lys phe 1eu arg ser glu ala val ala gln 1eu trp gly gln lys lys 1eu ser asn met thr tyr glu lys 1eu ser
991/331 1021/341 1051/351
CGG GCC ATG AGG TAC TAC TAC AAA CGG GAG ATC CTG GAA CGG GTG GAT GGC CGG CGA CTC GTC TAC AAG TTT GGC AAA AAC TCA AGC GGC
arg ala met arg tyr tyr tyr lys arg glu 1le 1eu glu arg val asp gly arg 1eu val tyr lys phe gly lys asn ser ser gly
1081/361 1111/371
TGG AAG GAG GAA GAG GTT CTC CAG AGT CGG AAC TGA
trp lys glu glu glu val 1eu gln ser arg asn OPA

Fig. 1 (cont.)

3/21

MAATCEISNIFSNYFSAMYSSEDSTLASVPPAATFGADDLVITLSNPQMSLEG 53

TEKASWILGEQPQFWSKTQVLDWISYQVEKNKYDASAIDFSRCDDMDGATLCNCA 106

LEELRLVFGPLGDQLHAQLRDLTSSSDELSWIILLEKDGMADFQEALDPGPF 159

DQGSPEAQELLDDGQQQASPYHPGSCGAGAPSPQSSDVSTAGTGASRSSSHSSDS 212

GGSDVDLDPTDGKLFPSDGTRDCKKGDPKHKGKRKGRPRKLSKEYWDCLEGKK 265

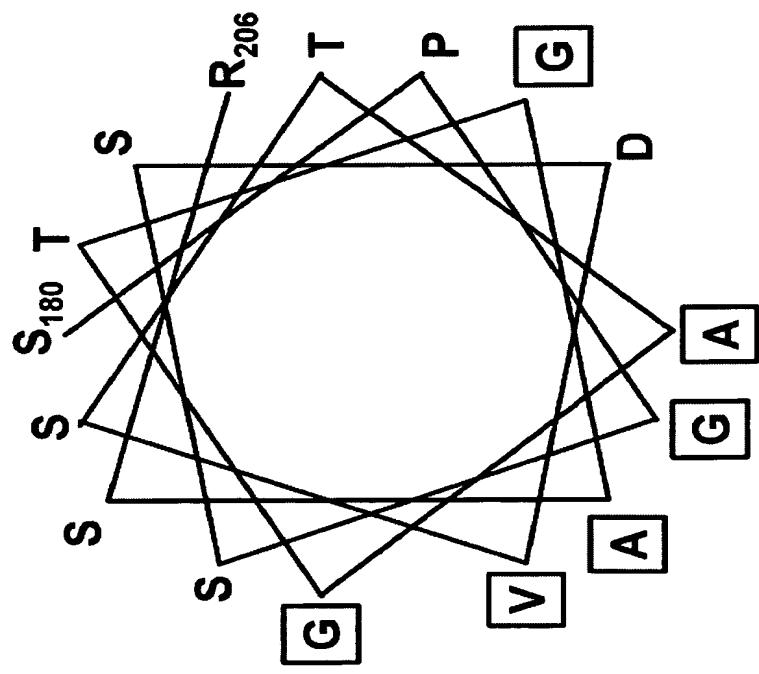
SKHAPRGTHIWEFIRDILIHPELNEGIMKWEINRHEGVFKFLRSEAVAQQLWGQK 318

KKNSNNMTYEKLSSRAMRYYYKREIELLERVDGRRILVYKF3KNSSGWKEEEEVLQSRN 371

Fig. 2A

Consensus+: P W V W E G LC
ESX (64-103) : PQFWWSKTQVLDWISYQVEKNKYDASAIDFSRCDDMDGATLC
P+ W++T V DW+ + V N++ +DF + M+GA LC
ETS-1 (69-106): PRQWTETHVRDWMMWV--NEFSLKGVDFQKFCMNGAAC

Fig. 2B



ESX (188-238) : APSPGSSSDVSTAGTGASSRSSHSSDGGSDVDLDPTDGKLFPSDGERDCKKG
 APS S S++ + S SS S SD + + I PS F
 SOX4 (370-420) : APSSAPSHASSSSSSSSSSSSSSSSSSSSDEFDDLLDINPSSNFESSMSLG

Fig. 2C

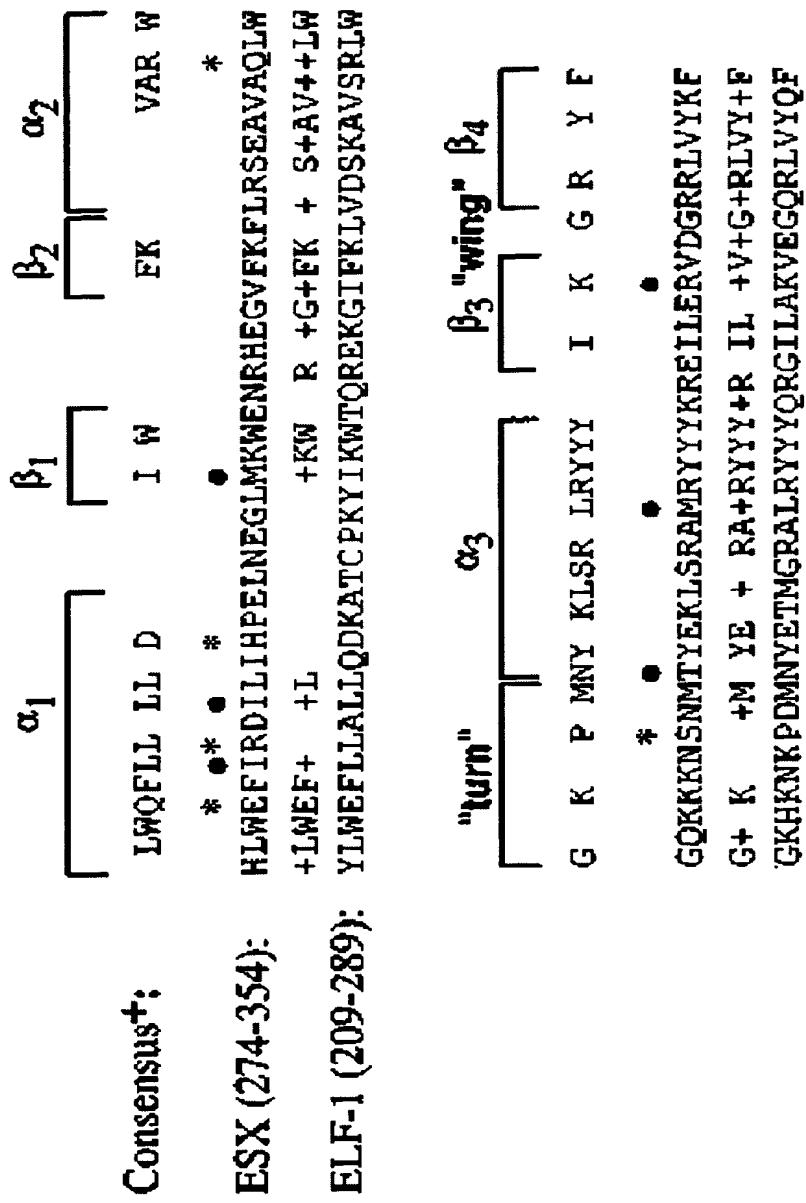


Fig. 2D

Human ESX Protein Sequence

MAATCEISNIFSNYFSAMYSSEDSTLASVPPAAATFGADDLVLTLSNPQMSLEG 53
TEKASWLGEPQQFWSKTQVLDWISYQVEKNKYDASAIDFSRCMDGATLCNCA 106
LEELRLVFGPLGQDQLHAQLRDLITSSSSDELSWIIELLEKDGMAFQEALDPGPF 159
DQGSPEAQLLDDGQQQSPYHPGSCGAGAPSPGSSSDVSTAGTGASRSSHSSDS 212
GGSDVDLDPTDGKLFPSDGFRDCKKGDPKHGKRKRGRRKLSKEYWDCLEGKK 265
SKHAPRGTHLWEFIRDILIHPELNEGIMKWENRHEGVFKFLRSEAVAQLWGQK 318
KKNSNMTYEKLSRAMYYKKREILLERVGDGRRLIVYKFGKNSSGWKEEEVLOSRN 371

Fig. 2E

7/21

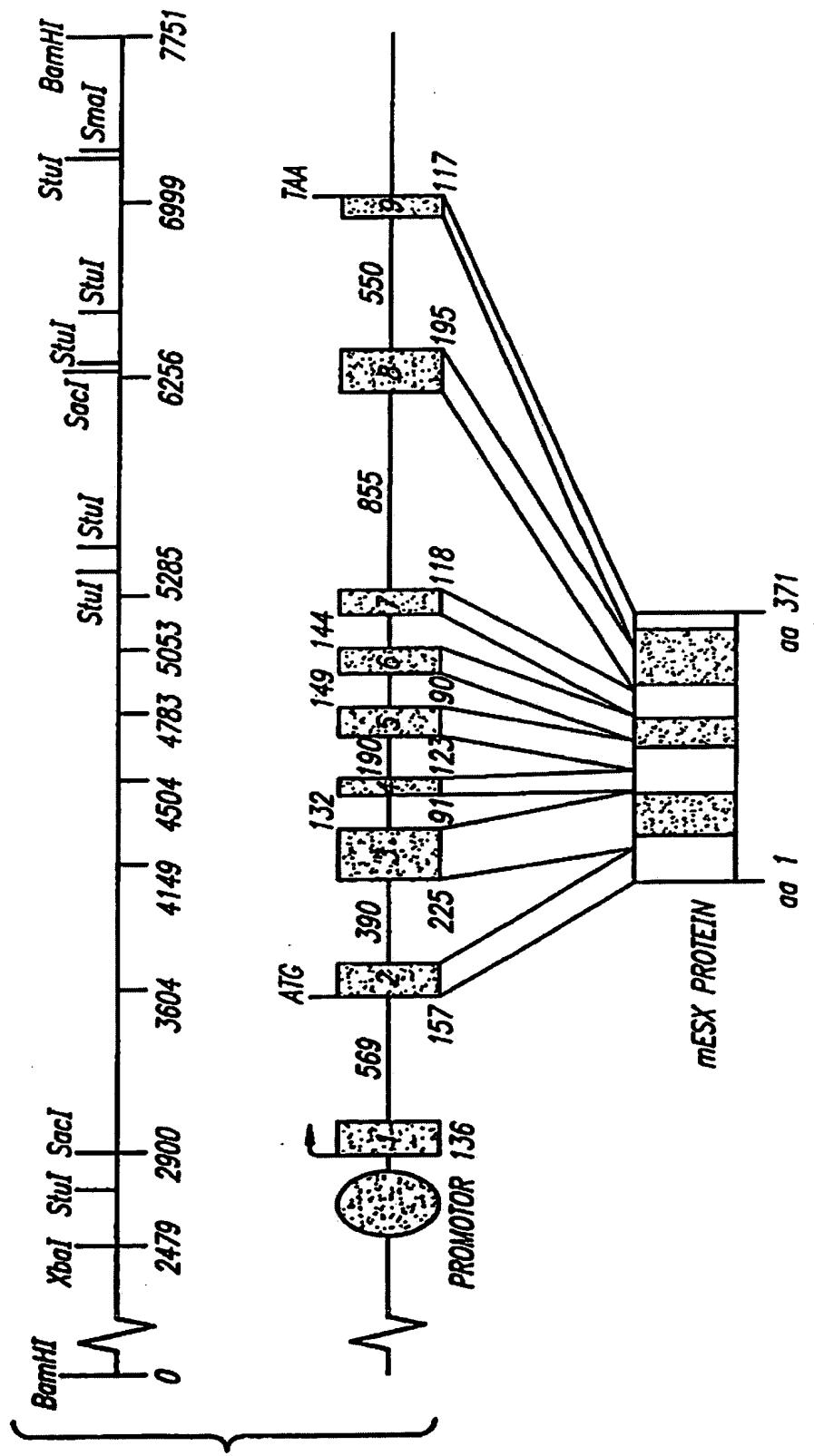


Fig. 3

00/21

1/1 31/11
ATG GCT GCA ACC TGT GAG ATT AGC AAC ATT TTT AGC AAC TAC TTC AGT GCG ATG TAC AGC
Met ala ala thr cys glu ile ser asn ile phe ser asn tyr phe ser ala met tyr ser
61/21 91/31

TCG GAG GAC TCC ACC CTG CCC TCT GTT CCC CCT GCT GCC ACC TTT GGG GGC GAT GAC TTG
ser glu asp ser thr leu ala ser val pro pro ala ala thr phe gly ala asp asp leu
121/41 151/51 E2 ▾ E3

GTA CTG ACC CTG AGC AAC CCC CAG ATG TCA TTG GAG GGT ACA GAG AAG GCC AGC TGG TTG
val leu thr leu ser asn pro gln met ser leu glu gly thr glu lys ala ser trp leu
181/61 211/71

GGG GAA CAG CCC CAG TTC TGG TCG AAG ACG CAG GTT CTG GAC TGG ATC AGC TAC CAA GTG
gly glu gln pro gln phe trp ser lys thr gln val leu asp trp ile ser tyr gln val
241/81 271/91

GAG AAC AAG TAC GAC GCA AGC GCC ATT GAC TTC TCA CGA TGT GAC ATG GAT GGC GCC
glu lys asn lys tyr asp ala ser ala ile asp phe ser arg cys asp met asp gly ala
301/101 331/111

ACC CTC TGC AAT TGT GCC CTT GAG GAG CTG CGT CTG GTC TTT GGG CCT CTG GGG GAC CAA
thr leu cys asn cys ala leu glu glu leu arg leu val phe gly pro leu gly asp gln
361/121 E3 ▾ E4 391/131

CTC CAT GCC CAG CTG CGA GAC CTC ACT TCC AGC TCT TCT GAT GAG CTC AGT TGG ATC ATT
leu his ala gln leu arg asp leu thr ser ser ser ser asp glu leu ser trp ile ile
421/141 451/151 E4 ▾ E5

GAG CTG CTG GAG AAG GAT GGC ATG GCC TTC CAG GAG GCC CTA GAC CCA GGG CCC TTT GAC
glu leu leu glu lys asp gly met ala phe gln glu ala leu asp pro gly pro phe asp

Fig. 4

9/21

481/161 CAG GGC AGC CCC TTT GCC CAG GAG CTG GAC GAC GGT CAA GCC AGC CCC TAC CAC
gln gly ser pro phe ala gln glu 1eu 1eu asp asp gly gln gln ala ser pro tyr his
541/181 CCC GGC AGC TGT GCG GCA GGA GCC Ccc TCC CCT GGC AGC TCT GAC GTC TCC ACC GCA GGG
pro gly ser cys gly ala pro ser pro gly ser ser asp val ser thr ala gly
601/201 ACT GGT GCT TCT CGG AGC TCC CAC TCC TCA GAC TCC EGT GGA AGT GAC GTG GAC CTG GAT
thr gly ala ser arg ser ser his ser ser asp ser gly gly ser asp val asp 1eu asp
661/221 CCC ACT GAT GGC AAG CTC TTC CCC AGC GAT GGT TTT CGT GAC TGC AAG AAG GGG GAT CCC
pro thr asp gly lys 1eu phe pro ser asp gly phe arg asp cys lys 1ys gly asp pro
721/241 AAG CAC GGG AAG CGG AAA CGA GGC CGG CCC CGA AAG CTG AGC AAA GAG TAC TGG GAC TGT
lys his gly lys arg gly arg pro arg lys 1eu ser 1ys glu tyr trp asp cys
781/261 CTC GAG GGC AAG AAG AGC AAG CAC GCG CCC AGA GGC ACC CAC CTG TGG GAG TTC ATC CGG
1eu glu gly lys ser lys his ala pro arg gly thr his 1eu trp glu phe ile arg
841/281 GAC ATC CTC ATC CAC CCG GAG CTC AAC GAG GGC CTC ATG AAG TGG GAG AAT CGG CAT GAA
asp ile 1eu 1le his pro glu 1eu asn glu gly 1eu met 1ys trp glu asn arg his glu
901/301 GGC TTC aag TTC CTG CGC TCC GAG GCT GTG GCC CAA CTA TGG GGC CAA AAG AAA AAG
gly val phe lys phe 1eu arg ser glu ala val ala gln 1eu trp gly gln 1ys lys 1ys

Fig. 4 (cont.)

10/21

961/321	AAC AGC AAC ATG ACC TAC GAG AAG CTG AGC CGG GCC ATG AGG TAC TAC AAA CGG GAG asn ser asn met thr tyr glu lys leu ser arg ala met arg tyr tyr tyr lys arg glu	991/331	E8↓ E9
1021/341	ATC CTG GAA CGG GTG GAT GGC CGG CGA CTC GTC TAC AAG TTT GGC AAA AAC TCA AGC GGC ile leu glu arg val asp gly arg arg leu val tyr lys phe gly lys asn ser ser gly	1051/351	
1081/361	TGG AAG GAG GAA GAG GTT CTC CAG AGT CGG AAC TGA trp lys glu glu glu val leu gln ser arg asn OPA	1111/371	

Fig. 4 (cont.)

mESX	1 MAATCEISNVFSNYFNAMYSSEDPTLAPAPP.	49
hESX	1 MAATCEISNIFSNYFSAMYSSEDSTLASVPPAATFGADDLVTLNPQMS	50
mESX	50 LEGPEKASWTSERPQFWSKTQYLEWISYQVEKNKYDASSIDFSRCNMDG	99
hESX	51 LEGTEKASWLGEQPOFWSKTQVLWISYQVEKNKYDASAIDFSRCDMG	100
mESX	100 TLCSCALEELRLVFGPLGDQLHAQLRDLTSNSSDEL SWIIELLEKDGM	149
hESX	101 TLCNCALEELRLVFGPLGDQLHAQLRDLTSSSSDEL SWIIELLEKDGM	150
mESX	150 QESLGDLGSPDQGSPFAQELLDDGRQASPYYCSTYGP	199
hESX	151 QEAQ.DPGPFDQGSPFAQELLDDGQQASPYPHPGSCGAGAPSPGSSDVSTA	199
mESX	200 GTATPQSSHASDSGGSDVDLDLTESKVFP	249
hESX	200 GTGASRSSHSSDSGGSDVDLDPDGKLFPSDGFRDOKKGDPKHGKRGR	249
mESX	250 PRKLSKEYWDCLEGKSKHAPRGTHLWEFIRDILIHPELNEG	299
hESX	250 PRKLSKEYWDCLEGKSKHAPRGTHLWEFIRDILIHPELNEG	299

Fig. 5

12/21

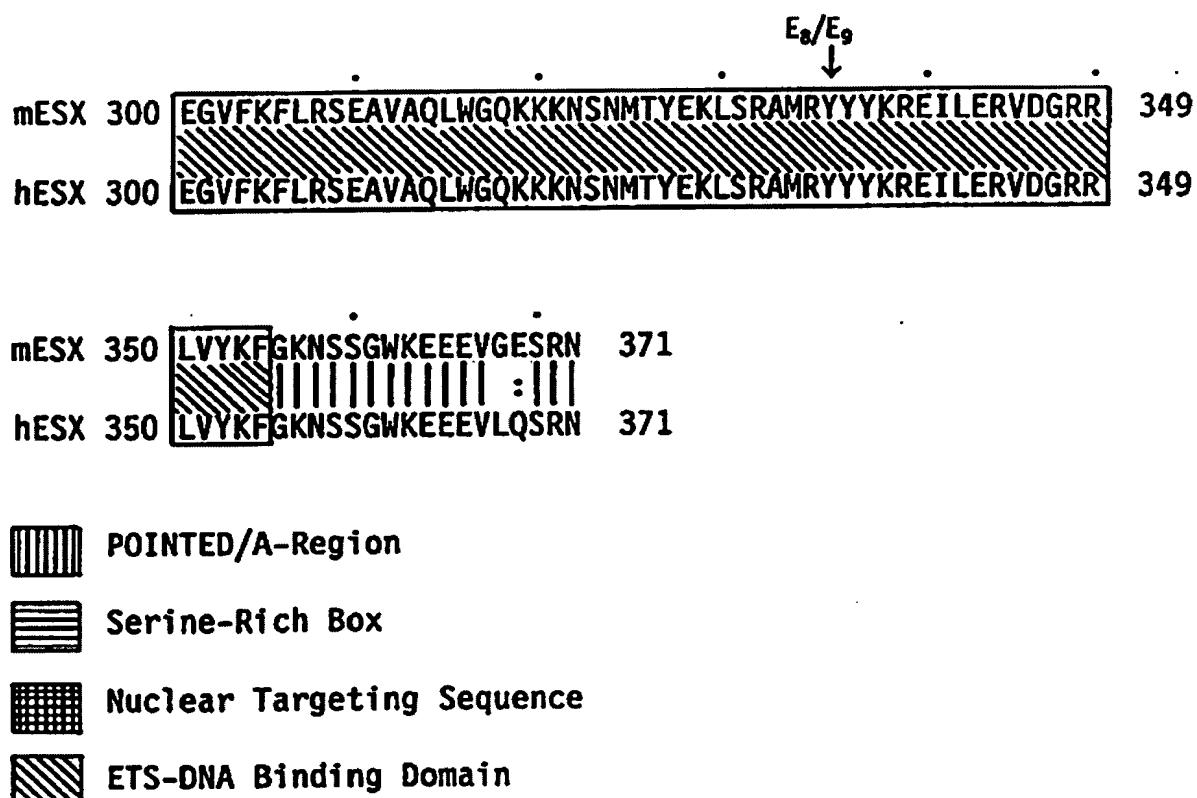


Fig. 5 (cont.)

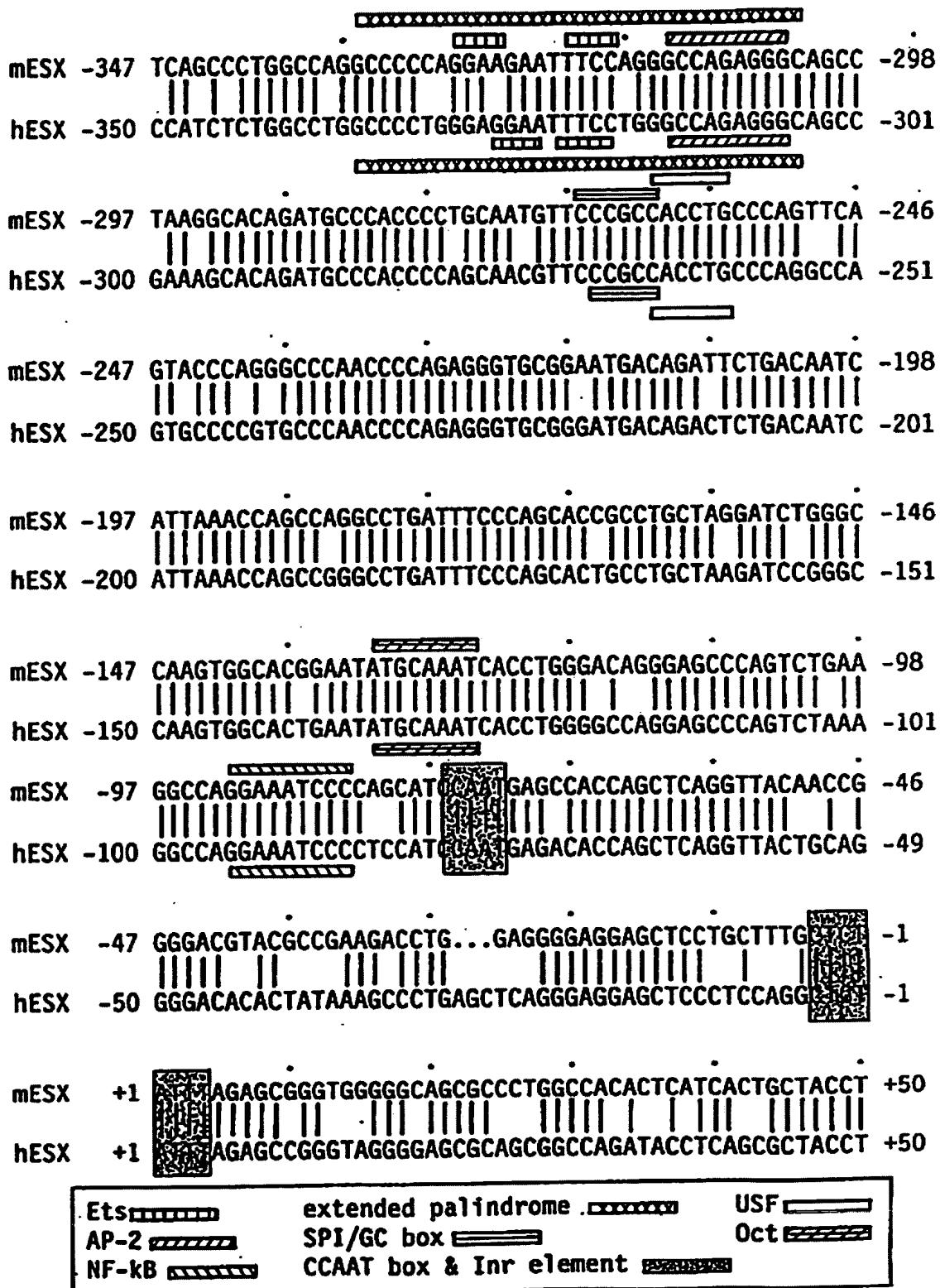


Fig. 6

14/21

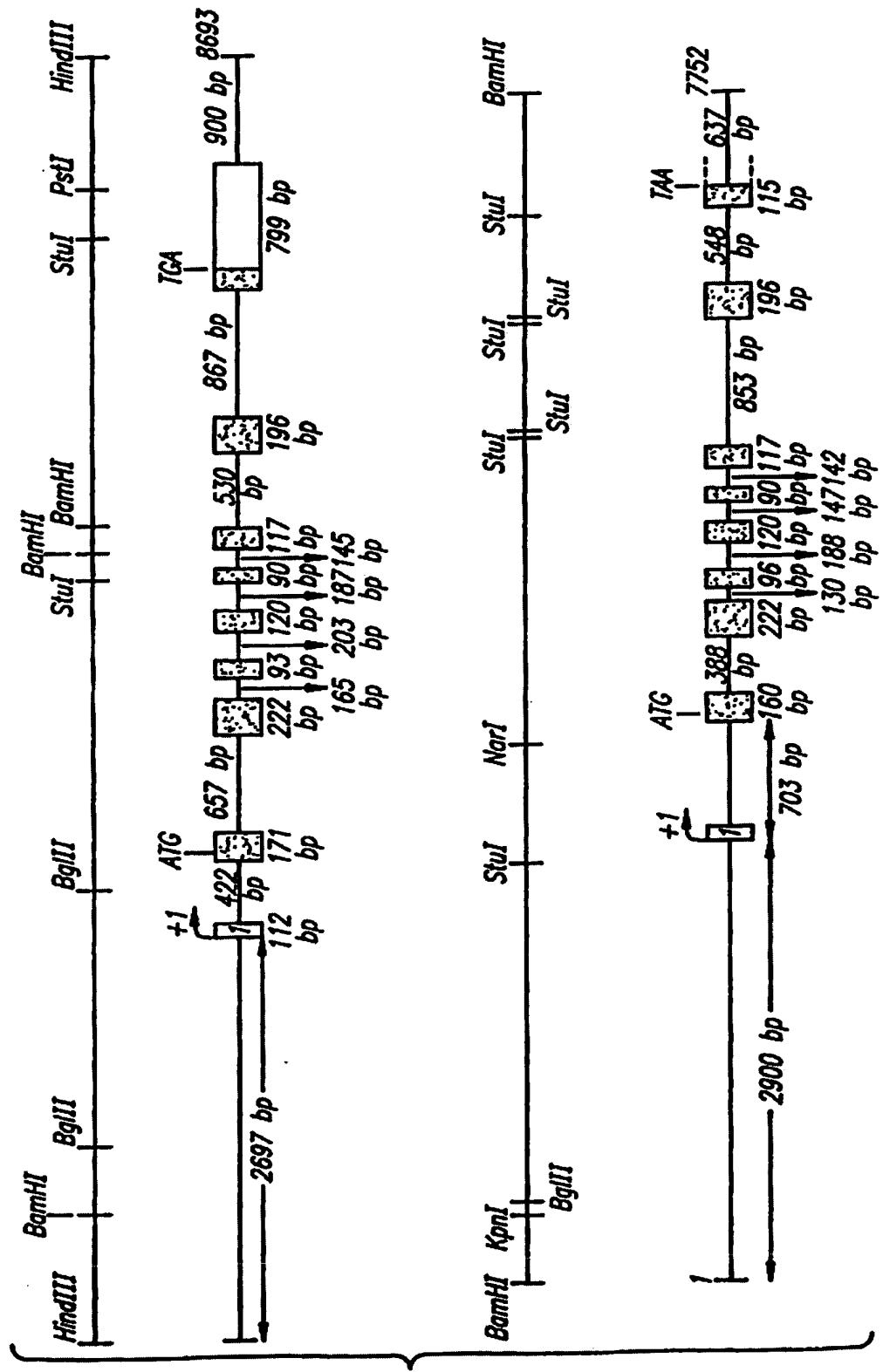
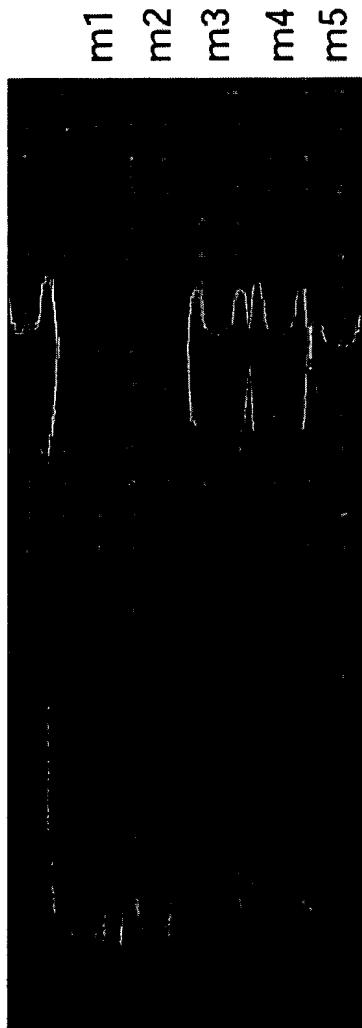


Fig. 7

15/21



WT 5' GGAGGAGGGCTGCTTGAGGAAGTATAAGAAT 3'
m1 5' ----- TA ----- 3'
m2 5' ----- C ----- 3'
m3 5' ----- AG ----- 3'
m4 5' ----- CC ----- 3'
m5 5' ----- C ----- 3'

Fig. 8A

16/21

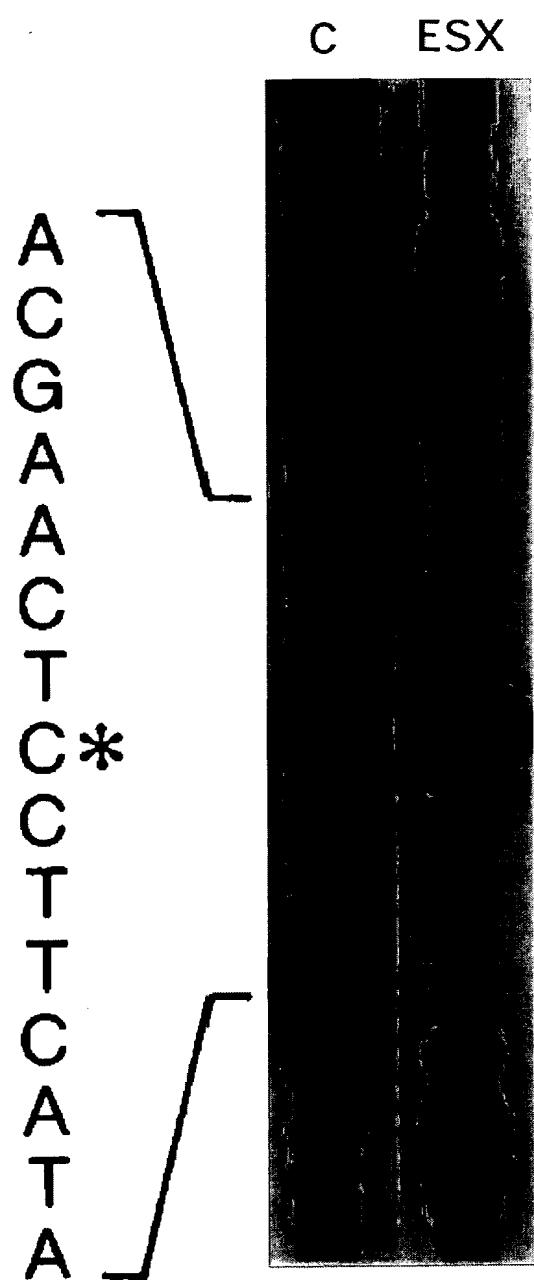


Fig. 8B

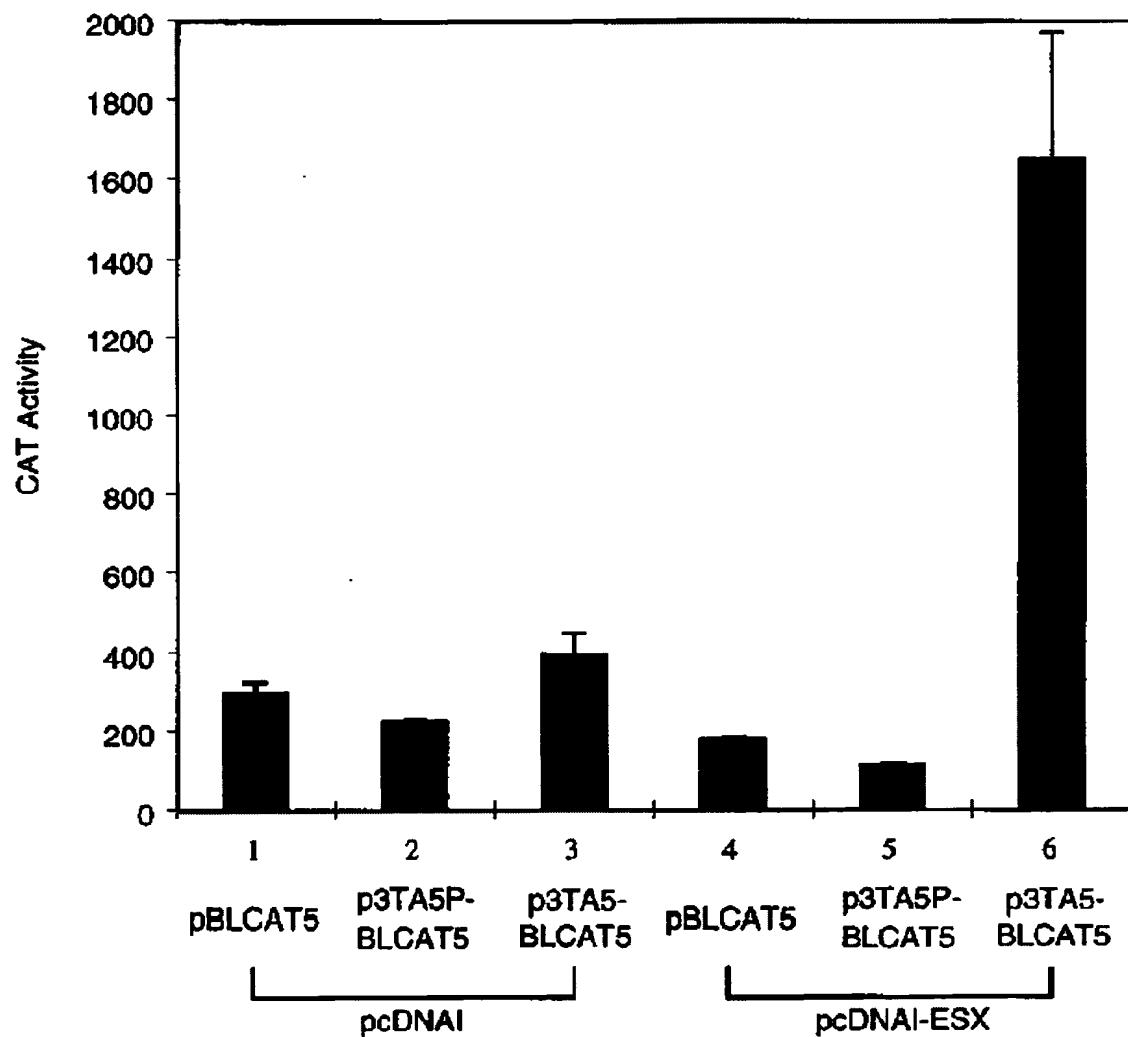


Fig. 8C

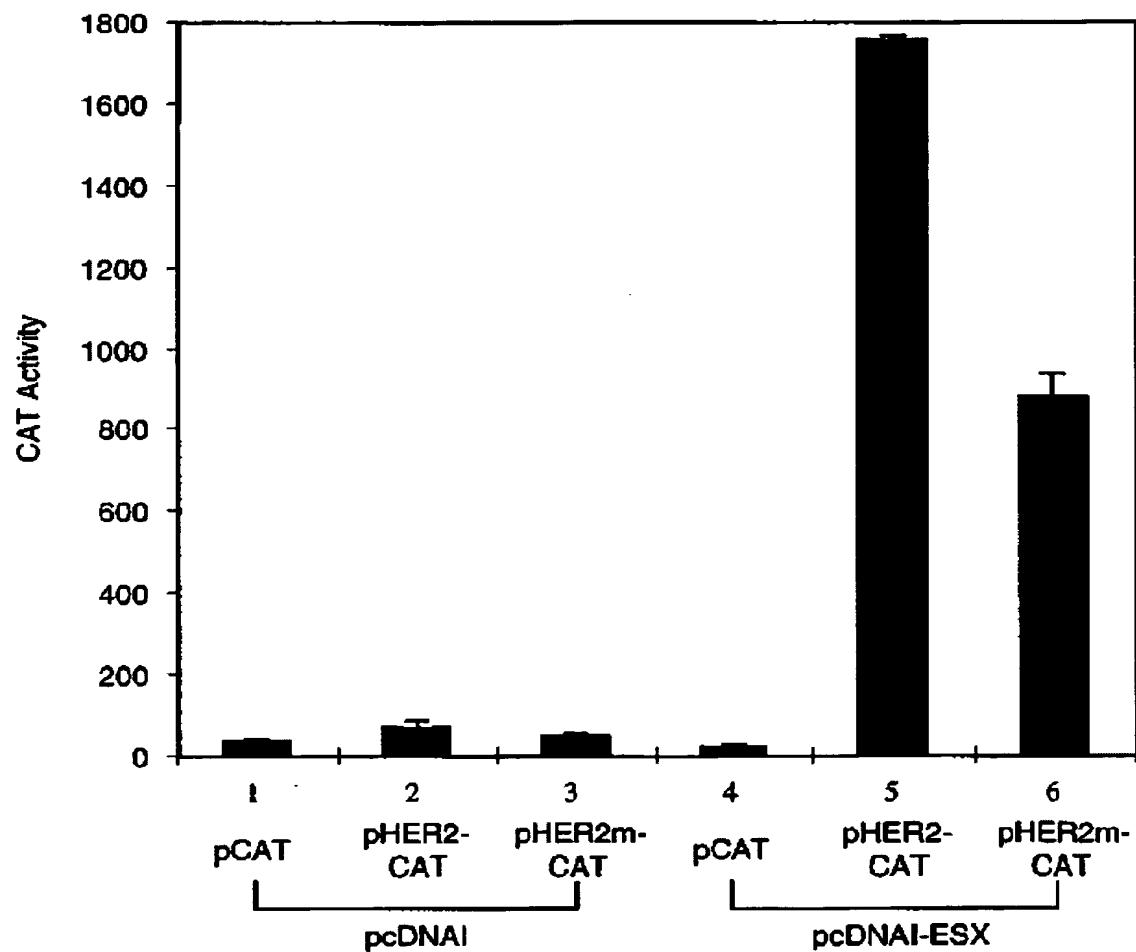


Fig. 8D

19/21

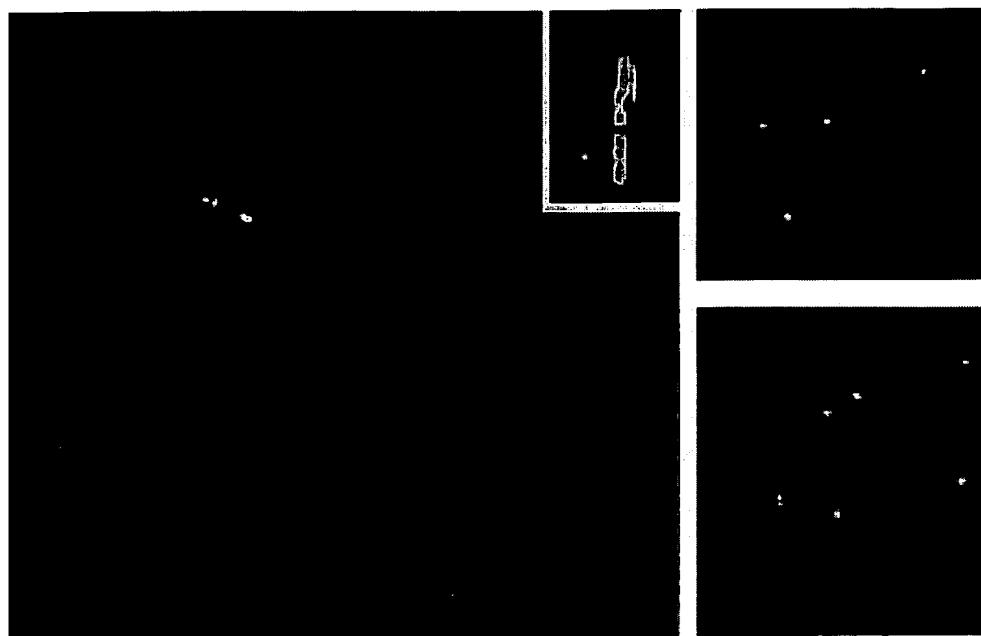


Fig. 8E

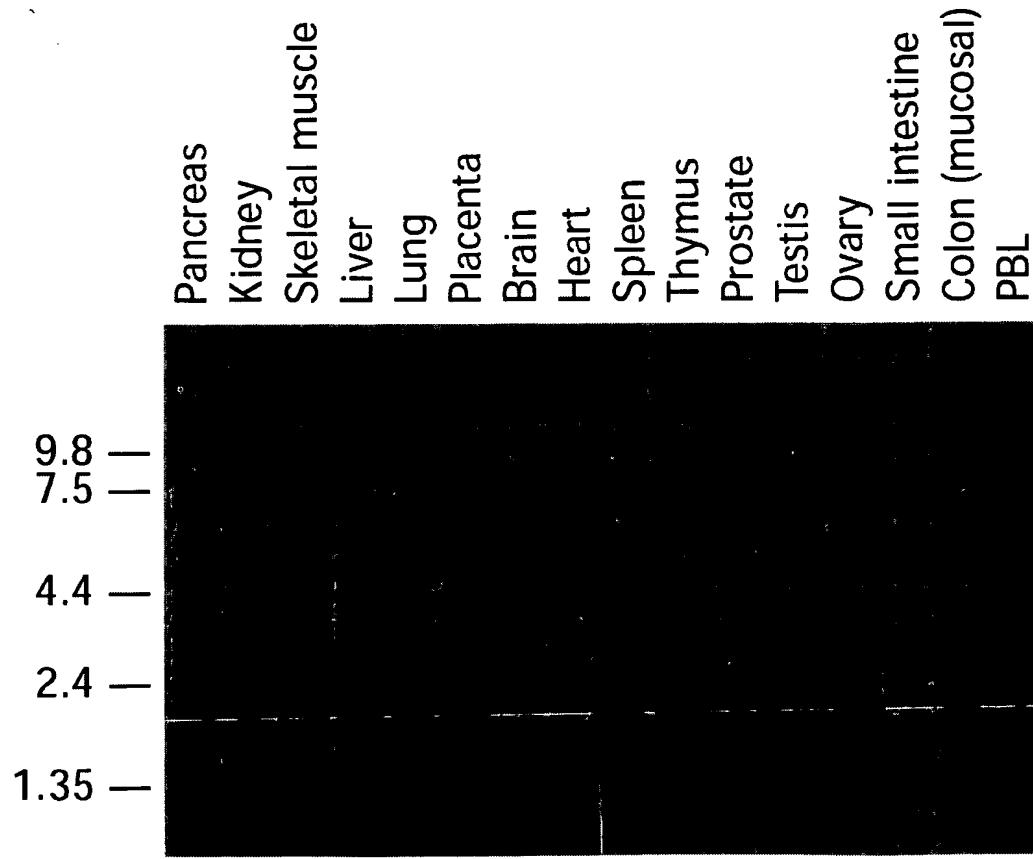


Fig. 9A

20/21

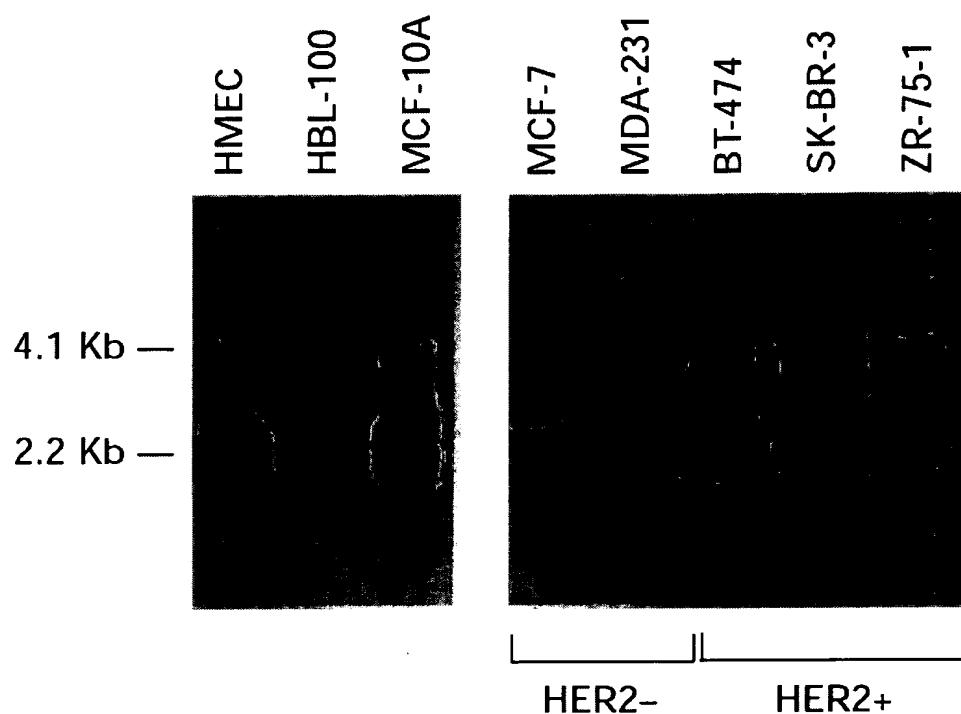


Fig. 9B

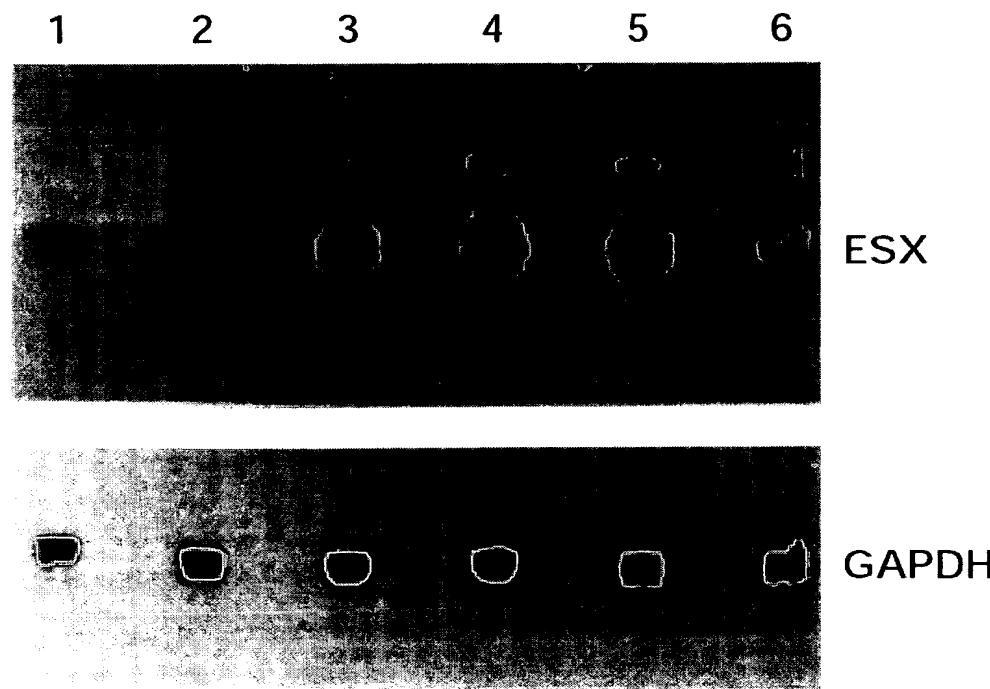


Fig. 9C

21/21

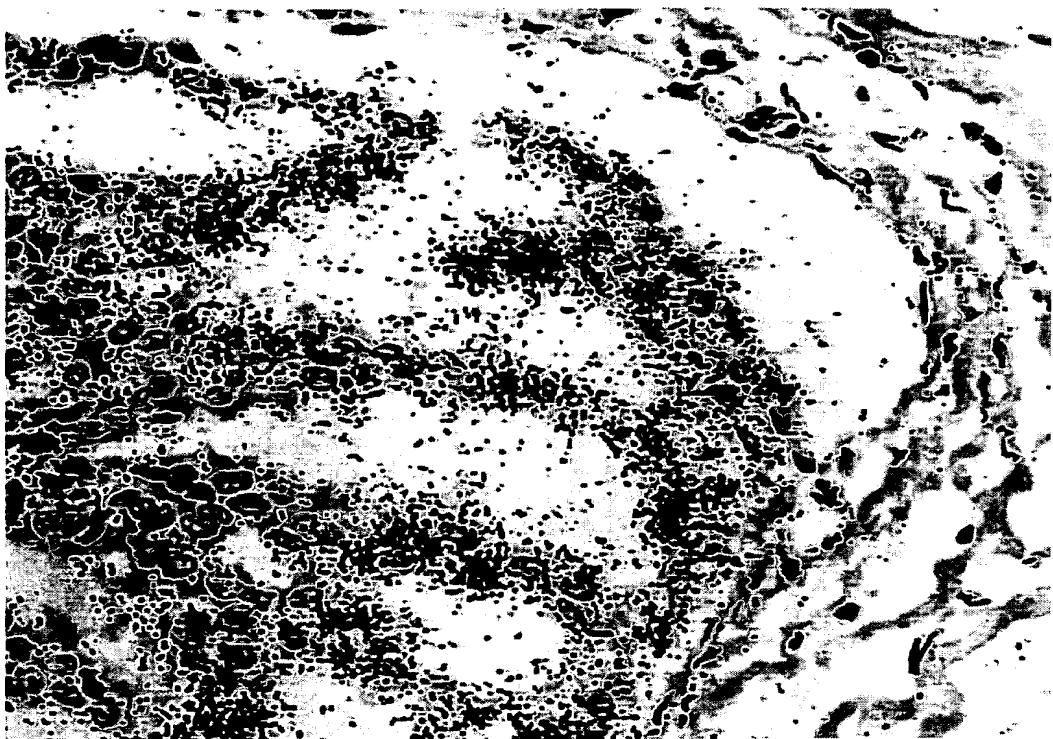


Fig. 10A



Fig. 10B